European Journal of Radiology 139 (2021) 109583

Contents lists available at ScienceDirect

European Journal of Radiology

journal homepage: www.elsevier.com/locate/ejrad

Research article

A multi-center study of COVID-19 patient prognosis using deep learning-based CT image analysis and electronic health records

Kuang Gong [a,1], Dufan Wu [a,1], Chiara Daniela Arru [a], Fatemeh Homayounieh [a], Nir Neumark [b], Jiahui Guan [c], Varun Buch [b], Kyungsang Kim [a], Bernardo Canedo Bizzo [b], Hui Ren [a], Won Young Tak [d], Soo Young Park [d], Yu Rim Lee [d], Min Kyu Kang [e], Jung Gil Park [e], Alessandro Carriero [f], Luca Saba [g], Mahsa Masjedi [h], Hamidreza Talari [h], Rosa Babaei [i], Hadi Karimi Mobin [i], Shadi Ebrahimian [a], Ning Guo [a], Subba R. Digumarthy [a], Ittai Dayan [b], Mannudeep K. Kalra [a,**], Quanzheng Li [a,*]

[a] Department of Radiology, Massachusetts General Hospital, Boston, United States
[b] MGH & BWH Center for Clinical Data Science, Boston, United States
[c] Nvidia, Boston, United States
[d] Department of Internal Medicine, School of Medicine, Kyungpook National University, Daegu, South Korea
[e] Department of Internal Medicine, Yeungnam University College of Medicine, Daegu, South Korea
[f] Radiologia, Azienda Ospedaliera Universitaria Maggiore della Carità, Novara, Italy
[g] Radiologia, Azienda Ospedaliera Universitaria Policlinico di Monserrato, Italy
[h] Department of Radiology, Kashan University of Medical Sciences, Kashan, Iran
[i] Department of Radiology, Firoozgar Hospital, Iran University of Medical Sciences, Tehran, Iran

ARTICLE INFO

Keywords:
COVID-19
Computed tomography
Deep learning
Electronic health records
Prognosis

ABSTRACT

Purpose: As of August 30th, there were in total 25.1 million confirmed cases and 845 thousand deaths caused by coronavirus disease of 2019 (COVID-19) worldwide. With overwhelming demands on medical resources, patient stratification based on their risks is essential. In this multi-center study, we built prognosis models to predict severity outcomes, combining patients' electronic health records (EHR), which included vital signs and laboratory data, with deep learning- and CT-based severity prediction.

Method: We first developed a CT segmentation network using datasets from multiple institutions worldwide. Two biomarkers were extracted from the CT images: total opacity ratio (TOR) and consolidation ratio (CR). After obtaining TOR and CR, further prognosis analysis was conducted on datasets from INSTITUTE-1, INSTITUTE-2 and INSTITUTE-3. For each data cohort, generalized linear model (GLM) was applied for prognosis prediction.

Results: For the deep learning model, the correlation coefficient of the network prediction and manual segmentation was 0.755, 0.919, and 0.824 for the three cohorts, respectively. The AUC (95 % CI) of the final prognosis models was 0.85(0.77,0.92), 0.93(0.87,0.98), and 0.86(0.75,0.94) for INSTITUTE-1, INSTITUTE-2 and INSTITUTE-3 cohorts, respectively. Either TOR or CR exist in all three final prognosis models. Age, white blood cell (WBC), and platelet (PLT) were chosen predictors in two cohorts. Oxygen saturation (SpO2) was a chosen predictor in one cohort.

Abbreviations: EHR, Electronic health records; COVID-19, Coronavirus disease of 2019; TOR, Total opacity ratio; CR, Consolidation ratio; GLM, Generalized linear model; WBC, White blood cell; PLT, Platelet; SpO2, Oxygen saturation; RT-PCR, Reverse-transcription polymerase chain reaction; MV, Mechanical ventilation; ICU, Intensive care unit; CT, Computed tomography; GGO, Ground-glass opacity; IRB, Institutional Review Board; GPU, Graphics processing unit; HU, Hounsfield unit; ESR, Erythrocyte sedimentation rate; AUC, Area under the curve; CI, Confidence interval; Hgb, Hemoglobin; MODS, Multiple Organ Dysfunction Score; SOFA, Sequential Organ Failure Assessment; LDH, Lactate dehydrogenase; hs-CRP, High-sensitivity C-reactive protein.

* Corresponding author at: 55 Fruit St, White 427, Boston, MA 02114, United States.
** Corresponding author at: 165 Cambridge Street, Boston, MA 02114, United States.
E-mail addresses: MKALRA@mgh.harvard.edu (M.K. Kalra), Li.Quanzheng@mgh.harvard.edu (Q. Li).
[1] indicates equal contributions.

https://doi.org/10.1016/j.ejrad.2021.109583
Received 2 September 2020; Received in revised form 28 January 2021; Accepted 1 February 2021
Available online 5 February 2021
0720-048X/

Conclusion: The developed deep learning method can segment lung infection regions. Prognosis results indicated that age, SpO2, CT biomarkers, PLT, and WBC were the most important prognostic predictors of COVID-19 in our prognosis model.

1. Introduction

Since its outbreak in November 2019, coronavirus disease of 2019 (COVID-19) has become a global pandemic due to its high contagiousness and lack of specific antiviral treatments or vaccines so far [1]. Reverse-transcription polymerase chain reaction (RT-PCR) [2,3] assay is currently the mainstay for the diagnosis of COVID-19 pneumonia. While early diagnosis is the key to initiate patient isolation and contact tracing to reduce its spread, assessment and prediction of the severity of COVID-19 pneumonia is important to initiate supportive treatment, project need for hospital admission and anticipate the use of mechanical ventilation (MV) and transition to intensive care unit (ICU). Computed tomography (CT) provides important information in COVID-19 pneumonia, especially for patients with moderate to severe disease as well as those with worsening cardiopulmonary status. CT helps evaluate the infection severity based on presence, extent, and type of pulmonary opacities such as ground-glass opacity (GGO), consolidation, and crazy-paving patterns [4–6]. Radiologists can grade the imaging severity of pulmonary opacities using a point-based scoring system [7,8]. However, such a scoring system is labor-intensive and can be time-consuming.

Recently deep learning has been widely applied to various medical imaging applications. As for COVID-19, there are several reports on utilizing deep learning for accurate differentiation of COVID-19 from other lung diseases and segmenting the opacity regions [9–16]. At the current resource-limited situation, prognosis prediction for COVID-19 is of paramount importance for patient management and decision making, e.g. MV usage and ICU admission. Several studies have explored the role of derived CT biomarkers for prognosis, alone [8,17] or combined with electronic health records (EHR) [18,19], all based on datasets from one institution or country. Considering the differences in medical resources and public health strategies [20,21], it is unclear if the risk factors identified would be the same for different regions.

In this work, we explored the risk factors for COVID-19 prognosis based on datasets from multiple institutions across the world. We first developed a deep learning-based CT segmentation network using datasets from different countries. CT-based biomarkers, total opacity ratio (TOR), and consolidation ratio (CR) were derived based on the network output. The CT biomarkers were used together with the EHR data for prognosis analysis to identify potential predictors for patient severity outcomes.

Contributions of this work include two aspects: (1) deep learning-based CT image segmentation for COVID-19 pneumonia was developed and validated based on datasets from different institutions; (2) COVID-19 patient prognosis analysis was conducted based on datasets from multiple cohorts using both CT-derived biomarkers and EHR data.

2. Materials and methods

2.1. Datasets

This study was approved by the respective Institutional Review Boards (IRBs) and informed consent forms were waived due to the retrospective nature of this study. The whole dataset consisted of 369 non-contrast chest CT examinations of patients with RT-PCR assay positive COVID-19 pneumonia scanned between January 1, 2020 and March 30, 2020, from five hospitals across the world. Of these, manual segmentation of all regions with pulmonary opacities related to COVID-19 pneumonia was available for 87/146 chest CTs from INSTITUTE-1, 25/102 chest CTs from INSTITUTE-2, 25/75 chest CTs from INSTITUTE-3, 8/8 chest CTs from INSTITUTE-4, and 14/14 chest CTs from INSTITUTE-5. The manually segmented CT datasets were used during network training and testing. Datasets from INSTITUTE-4 and INSTITUTE-5 do not contain EHRs, but only CT images, thus not included in the prognosis analysis but only CT segmentation-model development. For patients with multiple CT scans, we only included the first CT scan in this study.

The chest CT examinations were acquired on multidetector-row scanners (6–256 slices) from three CT vendors (GE Healthcare, Waukesha, US; Philips Healthcare, Eindhoven, Netherlands; Siemens Healthineers, Forchheim, Germany). All chest CT images were resampled to 5 mm section thickness and 256 × 256 resolution in the transverse plane. All datasets were initially graded for the type of pulmonary opacities by two fellowship-trained thoracic subspecialty radiologists (XX and XXX with 16 and 13 years of clinical experience in thoracic imaging), who supervised the annotation of each opacity by two post-doctoral research fellows (with 1–2 years of experience in chest CT research). These manually segmented opacity regions were used as ground truth during network training and testing. For the datasets with EHR, detailed summaries are presented in Table 1. A flowchart of the work is also given in Fig. 1 for better understanding of the data organization and processing.

2.2. CT image segmentation : neural network details

The network structure used in this work was based on a dense 3D network structure, which can fully extract features from input images because of better information propagation enabled by dense connections [22]. The Diagram of the network structure is shown in **Fig. A.1**. It is composed of an encoder path to extract contextual features and a decoder path to recover the image details, which contains the dense blocks as well as transition-down and transition-up modules. Each convolutional kernel has the size of 3 × 3 × 3. Transition-down module composes by a 1 × 1 × 1 convolution with stride 2 at X, Y axis which results in 1/2 spatial resolution reduction on the length and width of the input feature maps. It contains in-total 9 dense blocks, 4 transition-down and 4 transition-up modules. The last layer is a 1 × 1 × 1 convolution followed by the sigmoid activation function. The total number of trainable parameters for the network is 5.5 million.

To let the network focus on regions inside the lung, a lung mask was generated using a pretrained U-Net [23]. The network inputs were 3D CT images inside the lung mask and the outputs were the probability maps of the opacity regions inside the lung. The soft Dice [24], defined as

Table 1
Summary of the datasets with electronic health record (EHR) information from three different institutes, which were used for prognosis study in this work. SpO2 = peripheral oxygen saturation; WBC = white blood cell; PLT = platelet.

	INSTITUTE-1	INSTITUTE-2	INSTITUTE-3
Patients with EHR	146	102	58
Age (years)	60 ± 16 (21–100)	58 ± 18 (21–95)	68 ± 16 (30–95)
SpO2	91 ± 5 (66–100)	95 ± 6 (60–100)	96 ± 3 (86–100)
WBC	6551 ± 3177 (2100–27300)	6268 ± 2609 (2080–15000)	7915 ± 5253 (470–31670)
PLT	188,363 ± 69,617 (42000–453000)	240,765 ± 104,762 (82000–678000)	211,345 ± 91,156 (80000–587000)
MV/ICU/death	31	10	29
Death	21	6	18

$$SoftDice(\boldsymbol{p}, \boldsymbol{g}) = \frac{2\sum_{i}^{N} p_i g_i}{\left(\sum_{i}^{N} p_i^2 + \sum_{i}^{N} g_i^2\right)}, \quad \# \tag{1}$$

was utilized as the objective function, where $\boldsymbol{p}$ is the predicted probability map and $\boldsymbol{g}$ is the binary ground truth. The soft Dice is an approximation of the Dice coefficient, which can directly use the outputted probability map without pre-defining the threshold. During testing, we used a threshold of 0.5 to generate the binary segmentation mask to calculate the Dice coefficient. The network was implemented based on TensorFlow 1.14 [25]. The Adam optimizer [26] was used during network training to optimize the soft-dice objective function.

During training and testing, we used 3D patches as the network input due to the graphics processing unit (GPU) memory limit. The final patch size was 256 × 256 × 32. The sliding window was set to be 16 along the axial direction to generate the 3D patches. During training, the batch size was set to 1 due to the GPU memory limit. If a larger GPU memory is available, the batch size can be further increased to improve the network performance. Finally, one network was trained on the INSTITUTE-1 cohort (64 datasets for training, 23 datasets for validation, 77 datasets for testing), to generate the segmented regions of pulmonary opacities on the other cohorts. To obtain regions with abnormal pulmonary opacities on the INSTITUTE-1 cohort, another network was trained by combining other cohorts and then tested on INSTITUTE-1 (58 datasets for training, 14 datasets for validation, 87 datasets for testing). The training epoch number was 200.

2.3. CT image segmentation : biomarker derivations

The segmentation network output regions of pulmonary opacities (inclusive of all opacities such as ground glass, consolidation, mixed and crazy paving patterns). After obtaining the regions of pulmonary opacities, a threshold of -200 Hounsfield unit (HU) [27] was used to demarcate the consolidation regions inside each pulmonary opacity. Signs of consolidation regions can indicate advanced or more severe disease forms. Both the consolidation regions and the pulmonary-opacity regions are useful for CT biomarker derivations. In our study, two biomarkers were extracted from the CT images based on the network output: TOR, which was defined as the volume of pulmonary-opacity regions divided by the total lung volume; and CR, which was defined as the volume of consolidated regions divided by the total lung volume. For prognosis and survival analysis, both the TOR and CR derived from the CT images were tested as potential predictors. Note that all TOR and CR were calculated based on the network prediction and no TOR and CR were based on manual segmentations.

2.4. Prognosis analysis: data organization

After obtaining TOR and CR, further prognosis analysis was conducted on datasets from three institutions: INSTITUTE-1, INSTITUTE-2, and INSTITUTE-3. The EHR included patients' demographics (age), vital signs (temperature and peripheral oxygen saturation [SpO2]), and lab tests. The lab tests available from INSTITUTE-1 are platelet (PLT), white blood cell (WBC), lymphocyte percentage (Lymph%), erythrocyte sedimentation rate (ESR), lactate dehydrogenase (LDH) and C-reactive protein (CRP). The lab tests available from INSTITUTE-2 are PLT, WBC, Lymph%, LDH, hemoglobin (Hgb), procalcitonin (PCT), aspartate aminotransferase (AST), alanine aminotransferase (ALT), creatinine (Cr), estimated glomerular filtration rate (eGFR), sodium (Na), potassium (K), total bilirubin (TB), alkaline phosphatase (ALP), albumin (ALB) and blood urea nitrogen (BUN). The lab tests available from INSTITUTE-3 are PLT, WBC, Lymph%, LDH, Hgb, PCT, AST, ALT, Cr, eGFR, Na, K, CRP, Glucose and Neutrophil. Detailed EHR summaries are shown in Table 1. The correlation between TOR computed from CT images and EHR is shown in supplementary **Table A.1**. It should also be noted that the EHR data came from different timepoints: it represented the last measurement before the CT scan in INSTITUTE-3 but on-admission measurement in INSTITUTE-1 and INSTITUTE-2.

For both Institute 1 and Institute 2, the final adverse outcome events were ICU admission or death. For Institute 3, the final-outcome event was MV usage or death. In Institute 3, the concept of ICU admission was obscure as some medical units and post-operation suites were also utilized for intensive care due to ICU shortage. The clinical decision of mechanical ventilation usage in Institute 3 is similar to ICU admissions. Thus we used MV usage or death in Institute 3 as outcome events to

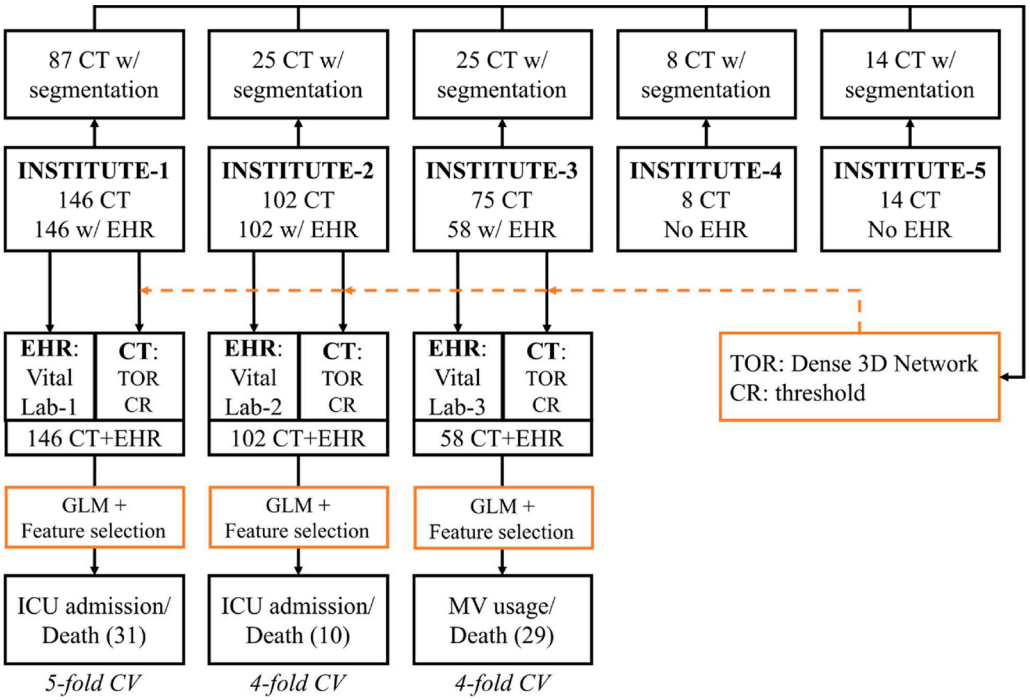

Fig. 1. Flowchart of the whole study. A dense 3D network was firstly trained with manual segmentation annotations as ground truth to predict the TOR. CR was then derived from the segmented results by thresholding at -200HU. The trained network was then used to process the CT images with EHR from INSTITUTE-1,2,3 to generate the corresponding TOR and CR biomarkers. At last, the TOR and CR biomarkers were combined with EHR. GLM and feature selection were used at INSTITUTE-1,2,3 independently to predict adverse outcome events as demonstrated in the figure. Cross validation (CV) was implemented using different number of folds at each site due to the different number of available datasets. Note that the three sites used the same vital sign variables but different lab tests. CR = consolidation ratio; TOR = total opacity ratio; GLM: generalized linear model; EHR: electronic health records; ICU: intensive care unit; CV: cross validation.

match with ICU admissions or death in Institutes 1 and 2.

2.5. Prognosis analysis: model construction

As for model construction, the generalized linear model (GLM) was used for the model analysis as the predictors employed in the analysis were continuous variables. The GLM model was constructed using the '*glm*' function in R. For each model, the optimal cutoff was calculated based on Youden's index (sensitivity + specificity - 1) and the corresponding sensitivity and specificity was obtained. The AUC values based on cross-validations were used to determine the variables ultimately used or discarded in the final models. AUC along with the 95 % AUC confidence interval (CI) were calculated based on 10,000 stratified bootstrap replicates. The *p*-value of the Z-statistics in each GLM model can be used to examine the relative importance of each chosen variable. For the dataset from each institute, we first used CT biomarkers, demographics and vital signs (e.g. age, patient's temperature, SpO2, TOR and CR) as potential predictors and obtained the model with the best AUC, which was considered as the basic model (model #1). Predictors from lab tests, e.g. platelet (PLT) count and white blood cell (WBC) count, were added to the basic model and the model with the best AUC was considered as the final model (model #2). To make full of the datasets, 5-fold, 4-fold and 4-fold stratified cross-validations were used for the INSTITUTE-1, INSTITUTE-2, and INSTITUTE-3 cohorts, respectively. The reason why different numbers of folds were used is due to the difference in total number of datasets from each institution.

For INSTITUTE-1 dataset, information regarding the adverse outcome (death or discharge) and the total in-hospitalization days are available, which enables further survival analysis. The plot of the Kaplan–Meier estimator was performed for the group with predicted ICU-or-death = 1 and the group with predicted ICU-or-death = 0 to show the survival probability difference, which can test the effectiveness of the constructed prognosis models. To find out which predictor is important for patients' survival, the Cox Proportional-Hazards regression model was used to evaluate the predictors (age, SpO2, WBC, PLT, Erythrocyte sedimentation rate (ESR), TOR, CR). The hazard ratio plots were calculated to indicate the importance of different predictors.

3. Results

3.1. CT image segmentation results

Fig. 2(**A, B, C**) shows three network segmentation examples, which demonstrates the feasibility of deep learning-based CT image segmentation. The mean Dice value was 0.641, 0.713 and 0.657 for datasets from INSTITUTE-1, INSTITUTE-2, and INSTITUTE-3, respectively. For each cohort, the predicted TOR versus the TOR calculated from annotation for all lung lobes are shown in Fig. 2**(B, C, D)**. The Spearman correlation coefficient was 0.755, 0.919, and 0.824 for INSTITUTE-1, INSTITUTE-2, and INSTITUTE-3, respectively. In clinical practice, TOR can be discretized to different categories to evaluate the severity of COVID. Here we discretized TOR into four categories (<5%, 5–25 %, 25–50 %, >50 %) and used multi-class accuracy (correct predictions / total predictions) to evaluate performance of the trained segmentation network. The final four-class prediction accuracy was 0.731, 0.832, 0.827 for INSTITUTE-1, INSTITUTE-2 and INSTITUTE-3, respectively. Results based on these quantitative measures demonstrated that the deep learning model generalized well across different sites.

Fig. 3 shows the boxplot of the predicted TOR for each lung lobe based on all 369 CT datasets. The left and right lower lobes had higher TOR compared to others, consistent with the expected pattern of COVID-19 on chest CT. Fig. 4 shows the relationship between the predicted TOR and time since symptom onset based on INSTITUTE-2. A drop in TOR after 10 days of symptom onset in all lung lobes was consistent with improving pneumonia. However, the lack of clear trends of TOR in the early phase of infection (0–10 days) may indicate variations and/or

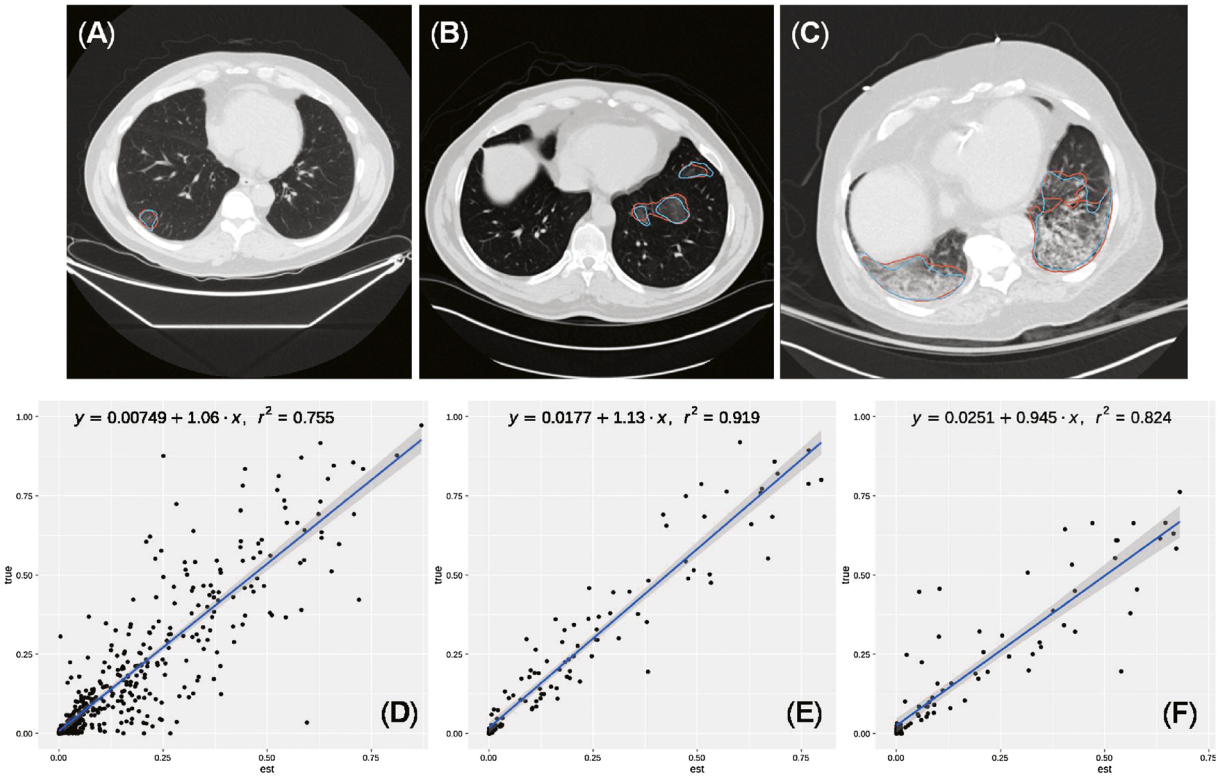

Fig. 2. Results of the CT segmentation network. (**A**,**B**,**C**) Three segmentation results overlayered on an axial slice. Blue curves indicate the manual segmentation contour, and red curves are the network predicted contour. (**D**, **E**, **F**) The predicted TOR versus the annotated TOR curves based on datasets from (**D**) INSTITUTE-1cohort, (**E**) INSTITUTE-2 cohort and (**F**) INSTITUTE-3 cohort. TOR = total opacity ratio.

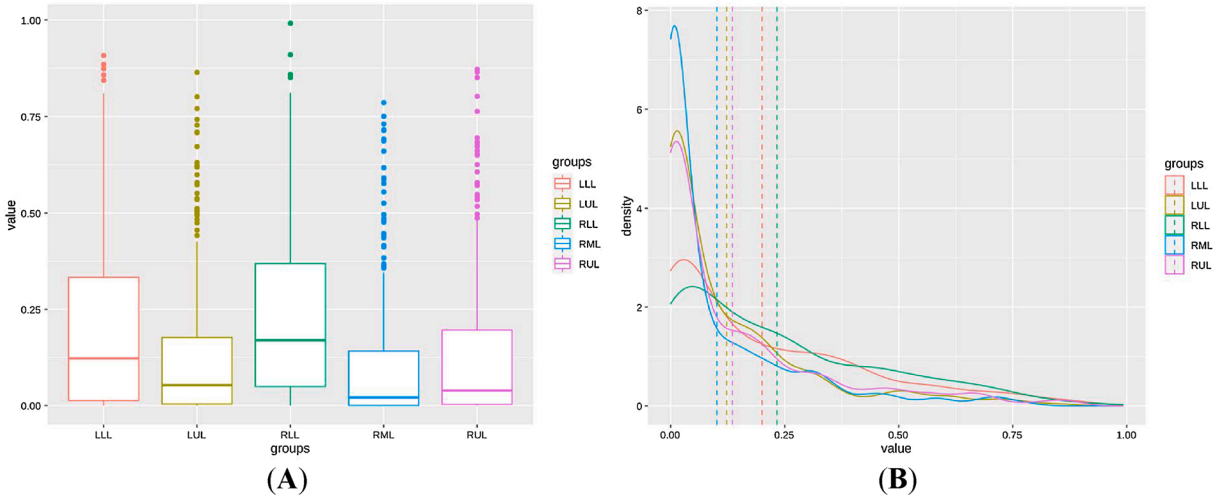

Fig. 3. **(A)** The boxplot quantification of all the lung lobes. X-axis stands for different lung lobes and y-axis stands for the pulmonary opacity ratio of the corresponding lung lobe. **(B)** The probability density of volume of pulmonary opacities for different lung lobes. X-axis stands for the pulmonary opacity ratio of the corresponding lung lobe and y-axis stands for the infection-ratio distribution probability density of the corresponding lung lobe. LUL = left upper lobe; LLL = left lower lobe; RUL = right upper lobe; RML = right middle lobe; RLL = right lower lobe; Whole = whole lung region.

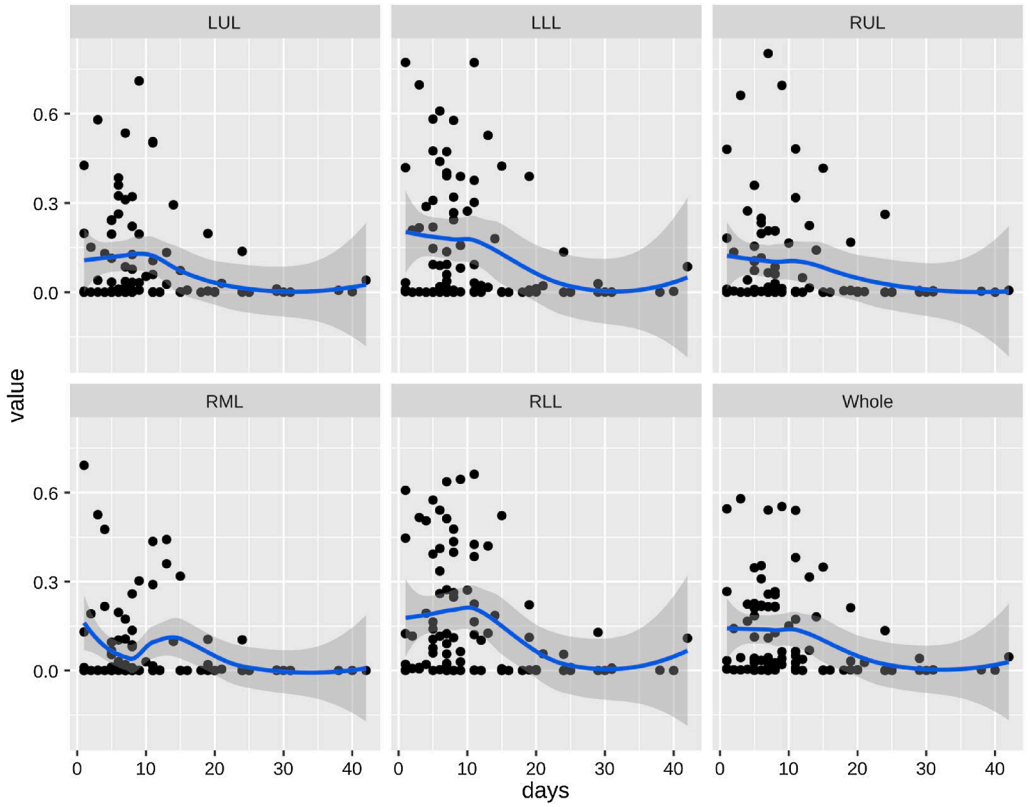

Fig. 4. Progression of regions with pulmonary opacities with symptom-onset days for the five lung lobes and the whole lung region. For each subplot, the x-axis stands for the days between symptom onset and the CT scan date and the y-axis stands for the pulmonary opacities ratio of the corresponding lung region. LUL = left upper lobe; LLL = left lower lobe; RUL = right upper lobe; RML = right middle lobe; RLL = right lower lobe; Whole = whole lung region.

unpredictability in disease progression within the assessed patient population.

3.2. Prognosis results

As for prognosis, ROC curves of model #1 and model #2 for INSTITUTE-1, INSTITUTE-2, and INSTITUTE-3 are shown in Fig. 5. For INSTITUTE-1. Optimal model #1 is based on age, SpO2 and CR. AUC(95 % CI) = 0.80(0.70, 0.88). The optimal cutoff is 0.16. The corresponding sensitivity and specificity is 0.81 and 0.68, respectively. Optimal model #2 is based on Age, SpO2, CR and PLT. AUC(95 % CI) = 0.85(0.77, 0.92). The optimal cutoff is 0.19. The corresponding sensitivity and specificity is 0.84 and 0.77, respectively. For INSTITUTE-2, optimal model #1 is based on CR. AUC(95 % CI) = 0.88(0.75, 0.97). The optimal cutoff is 0.05. The corresponding sensitivity and specificity is 0.90 and 0.83, respectively. Optimal model #2 is based on CR and WBC. AUC(95

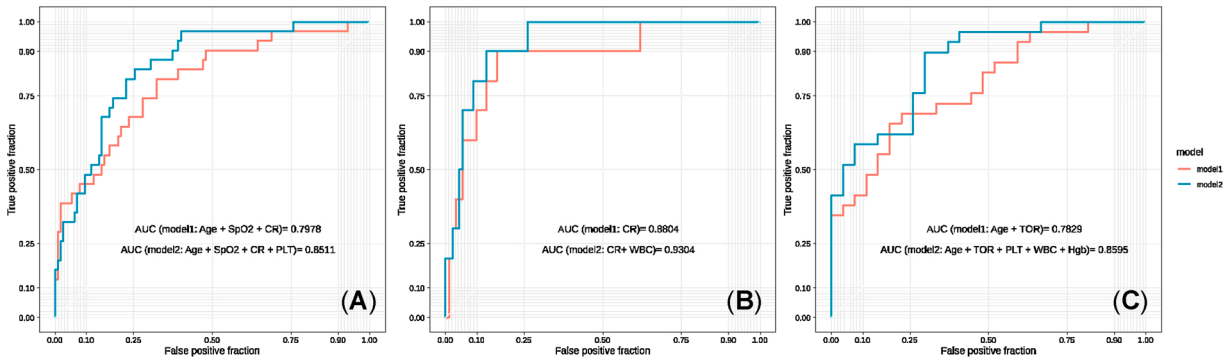

Fig. 5. ROC of the best predictors for the three cohorts: INSTITUTE-1, INSTITUTE-2, and INSTITUTE-3. Model #1 is the basic model constructed based on CT biomarkers, demographics, and vital signs. Model #2 is the final model constructed with data in Model #1 and additional lab test values. Detailed p-value statistics for each variable in each model are shown in **Table A.2** to check the importance of each variable in the prognosis model. (**A**) Results from cohort INSTITUTE-1. Optimal model #1 is based on age, SpO2 and CR. AUC(95 % CI) = 0.80(0.70, 0.88). Optimal model #2 is based on Age, SpO2, CR and PLT. AUC(95 % CI) = 0.85(0.77, 0.92). (**B**) Results from cohort INSTITUTE-2. Optimal model #1 is based on CR. AUC(95 % CI) = 0.88(0.75, 0.97). Optimal model #2 is based on CR and WBC. AUC(95 % CI) = 0.93(0.87, 0.98). (**C**) Results from cohort INSTITUTE-3. Optimal model #1 is based on Age and TOR. AUC(95 % CI) = 0.78(0.65, 0.89). Optimal model #2 is based on Age, TOR, PLT, WBC and Hgb. AUC(95 % CI) = 0.86(0.75, 0.94). CR = consolidation ratio; TOR = total opacity ratio; SpO2 = peripheral oxygen saturation; WBC = white blood cell; PLT = platelet; ESR = erythrocyte sedimentation rate; Hgb = hemoglobin; CI = confidence interval.

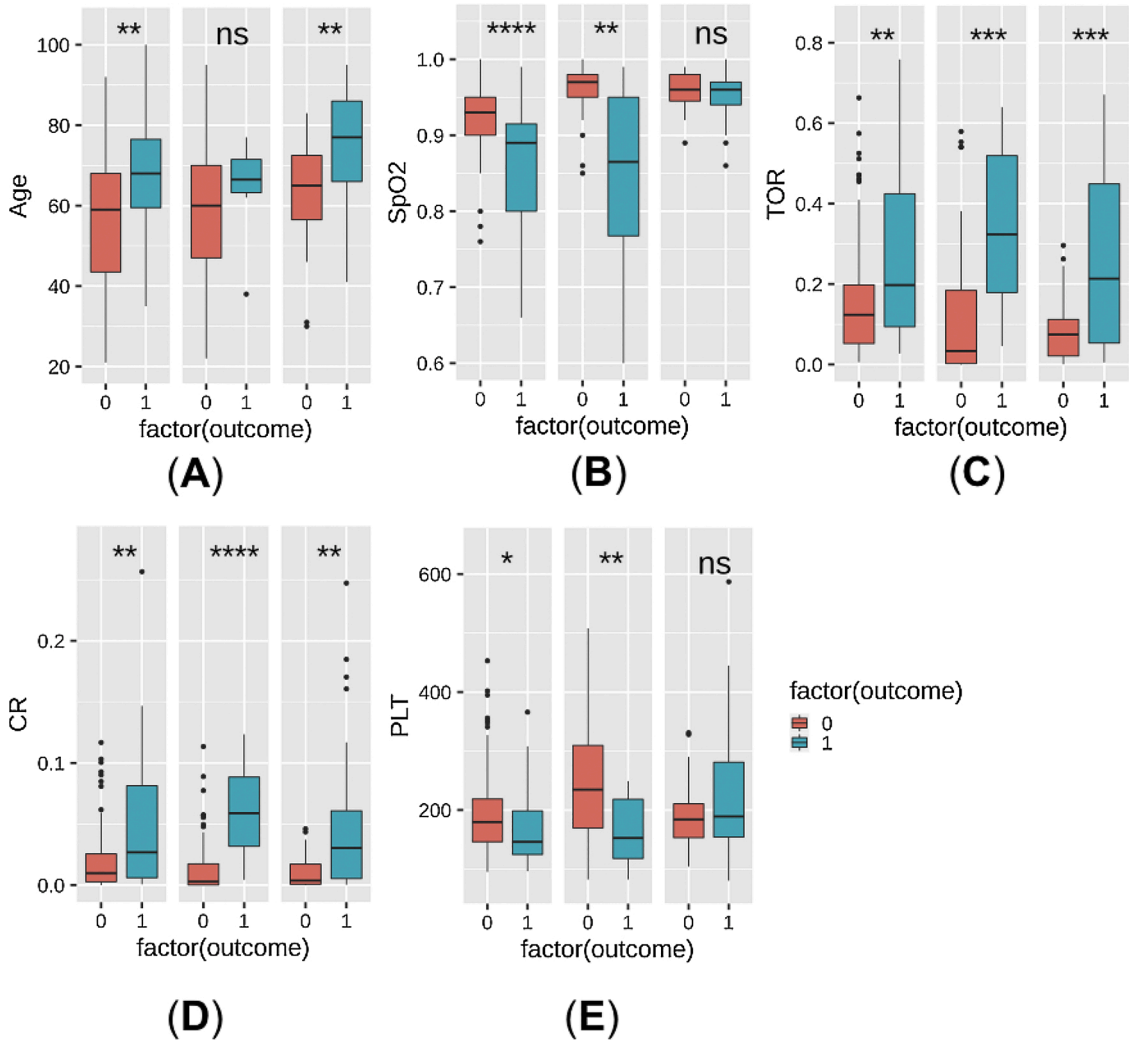

Fig. 6. Box and Whisker plot of (A) age; (B) SpO2, (C) TOR, (D) CR, and (E) WBC, in the three cohorts. For each sub-figure, the left, middle and right columns stand for the plot from the INSTITUTE-1cohort, INSTITUTE-2 cohort, and INSTITUTE-3 cohort, respectively. The red boxes are the patient group with non-severe outcomes and the blue boxes are patients with severe outcomes (MV/ICU/Death). The Wilcoxon signed-rank test was performed between the patient group with non-severe outcome and the patient group with severe outcomes. Correspondingly, *, **, and ns, located at the top of each bar plot, represents p-value<0.05, p-value <0.01, and non-significant, respectively. WBC = white blood cell; SpO2 = peripheral oxygen saturation; TOR = total opacity ratio; CR = consolidation ratio.

% CI) = 0.93(0.87, 0.98). The optimal cutoff is 0.07. The corresponding sensitivity and specificity is 0.90 and 0.87, respectively. For INSTITUTE-3, optimal model #1 is based on Age and TOR. AUC(95 % CI) = 0.78 (0.65, 0.89). The optimal cutoff is 0.58. The corresponding sensitivity and specificity is 0.66 and 0.81, respectively. Optimal model #2 is based on Age, TOR, PLT, WBC and Hgb. AUC(95 % CI) = 0.86(0.75, 0.94). The optimal cutoff is 0.35. The corresponding sensitivity and specificity is 0.90 and 0.70, respectively.

During cross-validation, each data fold will have one constructed GLM model. The *p*-values of the Z-statistics for each variable in each constructed GLM model are shown in supplementary **Table A.2**. For all three data cohorts, AUC $\geq$ 0.85 was achieved in the final models. For all three cohorts, the final models always had better AUC compared to the basic models, demonstrating the usefulness of lab tests in prognosis prediction. Either TOR or CR's existence in all three final models demonstrated CT's consistent prediction power of adverse disease outcomes for the assessed patient population. Among demographics features, age was an important component for both INSTITUTE-1 and INSTITUTE-3, but not a final predictor for INSTITUTE-2. It was also observed that SpO2 was an important factor in INSTITUTE-1, but not in the other two. Fig. 5 also indicates that WBC and PLT were chosen predictors in two of the three cohorts; hemoglobin (Hgb) was the chosen predictor in one of the two cohorts with this measure.

Fig. 6 summarizes the distribution of age, SpO2, TOR, CR, and WBC in patients with severe and non-severe disease in the three cohorts. No significant difference was found in patient age between the two groups in INSTITUTE-2 (*P-value* = 0.096). Neither it was found on SpO2 in INSTITUTE-3 (*P-value* = 0.471). This might explain why age and SpO2 were not predictors for model #1 in INSTITUTE-2 and INSTITUTE-3, respectively. As for the laboratory tests, no significant difference was found in INSTITUTE-1 regarding WBC (*P-value* = 0.341). This might explain why WBC was not a predictor for model #2 in INSTITUTE-1. From the correlation between TOR and EHR data shown in **Table A.1**, we can notice that TOR had a moderate negative correlation (-0.477) with SpO2 in INSTITUTE-2, which might explain why SpO2 was not a chosen predictor in INSTITUTE-2. Significant difference was found in TOR and CR for all three cohorts (*P-value* < 0.004), which might justify why CT derived biomarkers (TOR or CR) were predictors in both models #1 and #2.

We also studied the risk factors related to patients' survival time based on datasets from INSTITUTE-1 where the time points of mortality/discharge were available. Fig. 7**(A)** shows the plot of the Kaplan-Meier estimator for the group with predicted ICU-or-death = 1 and the group with predicted ICU-or-death = 0 based on the optimized model #2. This plot shows distinguishable survival probabilities for these two groups. Fig. 7**(B)** presents the hazard ratio plot of different risks based on the Cox proportional-hazards model. The most important risk factors of mortality were age (*P-value* = 0.003), SpO2 (*P-value* <0.001) and CR (*P-value* = 0.017). TOR and laboratory tests (such as WBC, PLT, and ESR) were not significant factors based on the hazard ratio plot.

4. Discussion

There are various studies focusing on opacity-region segmentation from the COVID-19 CT datasets. Huang et al. [12] have developed a CT segmentation network in order to perform serial quantitative CT assessment of COVID-19. To make the network training more efficient, Shan et al. have devised a human-in-the-loop strategy to reduce the manual labelling efforts [28]. Chaganti et al. [27] have designed a deep learning pipeline to perform semantic segmentation and various severity measures together. Fan et al. [29] has developed a semi-supervised approach to alleviate the shortage of labelled data. To overcome the pitfalls of noisy labels, Wang et al. [30] have developed a noise-robust segmentation network through the usage of a novel Dice loss and an adaptive self-ensembling training framework. Wu et al. [31] have developed a weakly supervised network based on hybrid labels to segment both opacity and consolidation regions from COVID-19 CT datasets. Amyar et al. [16] used the multi-task loss to improve the segmentation of infection regions in COVID-19 CT images. In this work, we developed a 3D dense network for CT image segmentation based on datasets from five institutes across the world. The dense network structure improved the segmentation performance over traditional UNet by reducing the number of training parameters [32]. Two biomarkers, characterizing the opacities and consolidations, were derived from the network output. In addition, we further investigated the role of CT-derived biomarkers and EHRs (vital signs and laboratory tests) for prognosis analysis based on datasets from three institutes of different countries, which were not conducted in previous CT-segmentation studies and are major contributions of this work. It was demonstrated that the biomarkers derived from the network segmentation imposed

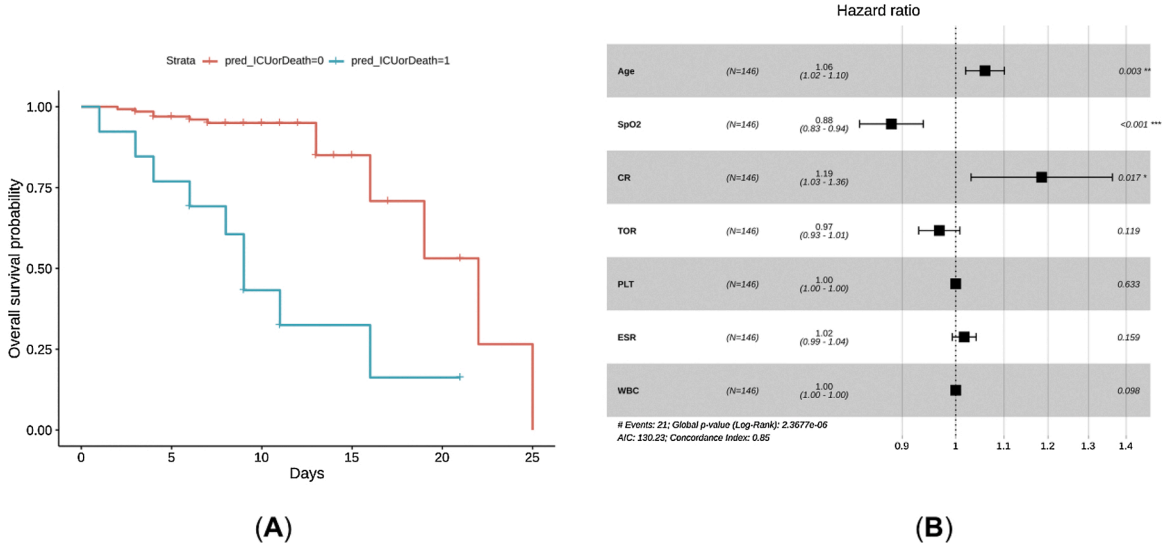

Fig. 7. Survival analysis based on the INSTITUTE-1 cohort. (**A**) The plot of the Kaplan-Meier estimator for the group with predicted-ICU-or-death = 1 and the group with predicted-ICU-or-death = 0. The prediction model is based on model #2 with AUC plot shown in Fig. 4(A). The cutoff threshold was set to 0.649, calculated based on R function 'optimalCutoff' with the default optimization criterion 'misclasserror'. (**B**) The hazard ratio plot regarding predictors: Age, SpO2, TOR, CR, PLT, ESR, and WBC. PLT = platelet; WBC = white blood cell; SpO2 = peripheral oxygen saturation; ESR = erythrocyte sedimentation rate; TOR = total opacity ratio; CR = consolidation ratio.

significant impact on predicting adverse outcomes for COVID-19 patients at three different sites (INSTITUTE-1,2,3), which indicated the prognostic values of CT-derived biomarkers for COVID-19 patients. In this work, we only used the soft Dice coefficient as the training loss to get the total infection regions and a simple thresholding method to calculate the consolidation regions. In the future, it can be combined with the aforementioned prior arts for more precise segmentation. In addition, for our datasets, when GGOs were manually segmented, the nodules were not considered as nodules only show up in small number of datasets. Thus our network cannot perform nodule segmentation. This is one limitation of our current datasets. Multi-class segmentation, including both GGO and nodules, deserves further investigations. We also noticed that the Dice coefficients from our CT segmentation network (0.641, 0.713 and 0.657) are lower compared to other segmentation tasks. Potential explanations for the low Dice coefficients is the variaty of manual labels by different radiologists and the small size of opacity regions in some datasets. If the prognosis model shows that the CT biomarkers derived from a simple network, which does achieve the best Dice coefficients, can work, it means CT derived biomarkers are useful for prognosis. Our prognosis results show that CT biomarkers are important for prognosis model construction. Thus, the relative low Dice coefficients do not invalidate our prognosis analysis.

Based on the prognosis analysis results in Figs. 5 and 7, CT derived biomarkers (TOR or CR), WBC and PLT were statistically significant predictors of adverse patient outcome. These findings are consistent with some recent risk-factor studies. Recent studies have suggested an association between CT severity scores generated from subjective evaluation of radiologists and mortality from COVID-19 pneumonia [8,17]. Prior publications suggested that a lower PLT count correlated with worse patient conditions such as higher Multiple Organ Dysfunction Score (MODS) [33,34] and Sequential Organ Failure Assessment (SOFA) score, known risk factors of COVID-19 mortality [35]. Abnormal WBC count is also observed in other COVID-19 datasets [35,38]. Previous studies [18,35–42] have also reported other laboratory tests as risk factors for adverse outcomes of COVID-19 pneumonia, including lactate dehydrogenase (LDH), lymphocyte count, and high-sensitivity C-reactive protein (hs-CRP). However, these factors were not significant predictors for prognosis based on our datasets. In [43], CT images and laboratory tests were combined for the prognosis analysis, which used datasets from only one country. It showed that combining CT and lab tests can help improve the prognostic accuracy, indicating Albumin and C-reactive protein as important risk factors. These laboratory tests were not present in our prognosis models based on our datasets. The difference between our prognosis model and the prior investigations stated above might be due to difference in treatments, sample sizes, measurement time, or population, which deserves further investigations.

Our current study has several limitations. Due to the small sample size of the ICU patients in INSTITUTE-2 (10 out of 102), it is difficult to conclude that age and SpO2 were not risk factors for these patients. In addition, the EHR data came from different timepoints: it represented the last measurement before the CT scan in INSTITUTE-3 but on-admission measurement in INSTITUTE-1 and INSTITUTE-2. The hospitalized patients were more likely to be on supplemental oxygen therapy in INSTITUTE-3, which can explain the insignificant difference of SpO2 between the two groups as shown in Fig. 6. We also want to point out that the percentage of CT opacity regions to the whole lung depends on the timepoint when the CT scan was performed. For our prognosis model, the timepoint when the CT scan was performed was not considered in our prognosis model, which is one limitation of our study and other studies which use CT opacity as a biomarker without considering the time course. Currently Chest X-ray is widely used for pre-screening and multiple timepoints of CXR are available. As the specificity of CXR is lower than CT regarding COVID-19, performance of biomarkers derived from CXR compared to CT is unknown. Conducting prognosis analysis based on CXR and EHRs is one of our future directions. Finally, in this work the prognosis model was constructed for each institute individually, without building a unified prognosis model for all institutes. Reasons for not building a united model is due to different treatment policies among the institutes and also the differences of measurement time. Pursuing a united prognosis model is one of our future works.

5. Conclusion

In conclusion, based on datasets from multiple institutions across the three continents, we have developed a generalizable deep learning model to segment opacity regions based on CT images from COVID-19 patients. Prognosis analysis was conducted using the derived CT biomarkers and EHR data. Results indicated that age, SpO2, CT-derived biomarkers, PLT, and WBC were the most important prognostic predictors of COVID-19 pneumonia in our prognosis models.

Ethical approval

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent

Informed consent form was waived due to the retrospective nature of this study.

CRediT authorship contribution statement

Kuang Gong: Conceptualization, Methodology, Software, Formal analysis, Writing - original draft, Writing - review & editing. **Dufan Wu:** Conceptualization, Methodology, Software, Formal analysis, Writing - original draft, Writing - review & editing. **Chiara Daniela Arru:** Software, Data curation, Writing - review & editing. **Fatemeh Homayounieh:** Software, Data curation, Writing - review & editing. **Nir Neumark:** Software, Writing - review & editing. **Jiahui Guan:** Software, Formal analysis, Writing - review & editing. **Varun Buch:** Software, Resources, Writing - review & editing. **Kyungsang Kim:** Formal analysis, Writing - review & editing. **Bernardo Canedo Bizzo:** Software, Resources, Writing - review & editing. **Hui Ren:** Formal analysis, Writing - review & editing. **Won Young Tak:** Resources, Writing - review & editing. **Soo Young Park:** Resources, Writing - review & editing. **Yu Rim Lee:** Resources, Writing - review & editing. **Min Kyu Kang:** Resources, Writing - review & editing. **Jung Gil Park:** Resources, Writing - review & editing. **Alessandro Carriero:** Resources, Writing - review & editing. **Luca Saba:** Resources, Writing - review & editing. **Mahsa Masjedi:** Resources, Writing - review & editing. **Hamidreza Talari:** Resources, Writing - review & editing. **Rosa Babaei:** Resources, Writing - review & editing. **Hadi Karimi Mobin:** Resources, Writing - review & editing. **Shadi Ebrahimian:** Resources. **Ning Guo:** Formal analysis, Writing - review & editing. **Subba R. Digumarthy:** Resources, Data curation, Writing - review & editing. **Ittai Dayan:** Resources, Writing - review & editing. **Mannudeep K. Kalra:** Conceptualization, Data curation, Writing - original draft, Writing - review & editing. **Quanzheng Li:** Conceptualization, Writing - original draft, Writing - review & editing.

Declaration of Competing Interest

The authors declare no conflict of interest.

Acknowledgements

This work was supported by the National Institutes of Health under grants RF1AG052653.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.ejrad.2021.109583.

References

[1] M. Cascella, M. Rajnik, A. Cuomo, S.C. Dulebohn, R. Di Napoli, Features, Evaluation and Treatment Coronavirus (COVID-19), StatPearls Publishing, 2020 (accessed May 17, 2020), http://www.ncbi.nlm.nih.gov/pubmed/32150360.

[2] T. Ai, Z. Yang, H. Hou, C. Zhan, C. Chen, W. Lv, Q. Tao, Z. Sun, L. Xia, Correlation of Chest CT and RT-PCR Testing in Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases, Radiology (2020) 200642, https://doi.org/10.1148/radiol.2020200642.

[3] Y. Fang, H. Zhang, J. Xie, M. Lin, L. Ying, P. Pang, W. Ji, Sensitivity of chest CT for COVID-19: comparison to RT-PCR, Radiology (2020) 200432, https://doi.org/10.1148/radiol.2020200432.

[4] A. Bernheim, X. Mei, M. Huang, Y. Yang, Z.A. Fayad, N. Zhang, K. Diao, B. Lin, X. Zhu, K. Li, S. Li, H. Shan, A. Jacobi, M. Chung, Chest CT findings in coronavirus Disease-19 (COVID-19): relationship to duration of infection, Radiology (2020) 200463, https://doi.org/10.1148/radiol.2020200463.

[5] W. Zhao, Z. Zhong, X. Xie, Q. Yu, J. Liu, Relation between chest CT findings and clinical conditions of coronavirus disease (COVID-19) pneumonia: a multicenter study, Am. J. Roentgenol. 214 (2020) 1072–1077, https://doi.org/10.2214/AJR.20.22976.

[6] F. Pan, T. Ye, P. Sun, S. Gui, B. Liang, L. Li, D. Zheng, J. Wang, R.L. Hesketh, L. Yang, C. Zheng, Time course of lung changes on chest CT during recovery from 2019 novel coronavirus (COVID-19) pneumonia, Radiology. (2020) 200370, https://doi.org/10.1148/radiol.2020200370.

[7] F. Feng, Y. Jiang, M. Yuan, J. Shen, H. Yin, D. Geng, J. Xu, Y. Hua, J. Shi, Y. Shi, Z. Zhang, Association of radiologic findings with mortality in patients with avian influenza H7N9 pneumonia, PLoS One 9 (2014), https://doi.org/10.1371/journal.pone.0093885.

[8] M. Yuan, W. Yin, Z. Tao, W. Tan, Y. Hu, Association of radiologic findings with mortality of patients infected with 2019 novel coronavirus in Wuhan, China, PLoS One 15 (2020) e0230548, https://doi.org/10.1371/journal.pone.0230548.

[9] L. Li, L. Qin, Z. Xu, Y. Yin, X. Wang, B. Kong, J. Bai, Y. Lu, Z. Fang, Q. Song, K. Cao, D. Liu, G. Wang, Q. Xu, X. Fang, S. Zhang, J. Xia, J. Xia, Artificial Intelligence Distinguishes COVID-19 From Community Acquired Pneumonia on Chest CT, 2020, https://doi.org/10.1148/radiol.2020200905.

[10] C. Butt, J. Gill, D. Chun, B.A. Babu, Deep learning system to screen coronavirus disease 2019 pneumonia, Appl. Intell. (2020), https://doi.org/10.1007/s10489-020-01714-3.

[11] J. Wang, Y. Bao, Y. Wen, H. Lu, H. Luo, Y. Xiang, X. Li, C. Liu, D. Qian, Prior-attention residual learning for more discriminative COVID-19 screening in CT images, IEEE Trans. Med. Imaging (2020) 1, https://doi.org/10.1109/TMI.2020.2994908.

[12] L. Huang, R. Han, T. Ai, P. Yu, H. Kang, Q. Tao, L. Xia, Serial quantitative chest CT assessment of COVID-19: deep-learning approach, Radiol. Cardiothorac. Imaging. 2 (2020) e200075, https://doi.org/10.1148/ryct.2020200075.

[13] H. Kang, L. Xia, F. Yan, Z. Wan, F. Shi, H. Yuan, H. Jiang, D. Wu, H. Sui, C. Zhang, D. Shen, Diagnosis of coronavirus disease 2019 (COVID-19) with structured latent multi-view representation learning, IEEE Trans. Med. Imaging (2020), https://doi.org/10.1109/TMI.2020.2992546.

[14] Y. Cao, Z. Xu, J. Feng, C. Jin, X. Han, H. Wu, H. Shi, Longitudinal assessment of COVID-19 using a deep learning–based quantitative CT pipeline: illustration of two cases, Radiol. Cardiothorac. Imaging. 2 (2020) e200082, https://doi.org/10.1148/ryct.2020200082.

[15] A. Abbasian Ardakani, U.R. Acharya, S. Habibollahi, A. Mohammadi, COVIDiag: a clinical CAD system to diagnose COVID-19 pneumonia based on CT findings, Eur. Radiol. (2020) 1–10, https://doi.org/10.1007/s00330-020-07087-y.

[16] A. Amyar, R. Modzelewski, H. Li, S. Ruan, Multi-task deep learning based CT imaging analysis for COVID-19 pneumonia: classification and segmentation, Comput. Biol. Med. 126 (2020) 104037, https://doi.org/10.1016/j.compbiomed.2020.104037.

[17] X. Qi, Z. Jiang, Q. YU, C. Shao, H. Zhang, H. Yue, B. Ma, Y. Wang, C. Liu, X. Meng, S. Huang, J. Wang, D. Xu, J. Lei, G. Xie, H. Huang, J. Yang, J. Ji, H. Pan, S. Zou, S. Ju, Machine learning-based CT radiomics model for predicting hospital stay in patients with pneumonia associated with SARS-CoV-2 infection: A multicenter study, MedRxiv (2020), https://doi.org/10.1101/2020.02.29.20029603, 2020.02.29.20029603.

[18] X. Bai, C. Fang, Y. Zhou, S. Bai, Z. Liu, L. Xia, Q. Chen, Y. Xu, T. Xia, S. Gong, X. Xie, D. Song, R. Du, C. Zhou, C. Chen, D. Nie, L. Qin, W. Chen, Predicting COVID-19 malignant progression with AI techniques, SSRN Electron. J. (2020), https://doi.org/10.2139/ssrn.3557984.

[19] H. Burdick, C. Lam, S. Mataraso, A. Siefkas, G. Braden, R.P. Dellinger, A. McCoy, J. L. Vincent, A. Green-Saxena, G. Barnes, J. Hoffman, J. Calvert, E. Pellegrini, R. Das, Prediction of respiratory decompensation in Covid-19 patients using machine learning: the READY trial, Comput. Biol. Med. 124 (2020) 103949, https://doi.org/10.1016/j.compbiomed.2020.103949.

[20] Y. Ji, Z. Ma, M.P. Peppelenbosch, Q. Pan, Potential association between COVID-19 mortality and health-care resource availability, Lancet Glob. Heal. 8 (2020) e480, https://doi.org/10.1016/S2214-109X(20)30068-1.

[21] J.S. Weitz, S.J. Beckett, A.R. Coenen, D. Demory, M. Dominguez-Mirazo, J. Dushoff, C.-Y. Leung, G. Li, A. Mägälie, S.W. Park, R. Rodriguez-Gonzalez, S. Shivam, C.Y. Zhao, Modeling shield immunity to reduce COVID-19 epidemic spread, Nat. Med. (2020) 1–6, https://doi.org/10.1038/s41591-020-0895-3.

[22] G. Huang, Z. Liu, L. Van Der Maaten, K.Q. Weinberger, Densely Connected Convolutional Networks, 2017, https://doi.org/10.1109/CVPR.2017.243.

[23] J. Hofmanninger, F. Prayer, J. Pan, S. Rohrich, H. Prosch, G. Langs, Automatic Lung Segmentation in Routine Imaging Is a Data Diversity Problem, Not a Methodology Problem, 2020. http://arxiv.org/abs/2001.11767.

[24] F. Milletari, N. Navab, S.A. Ahmadi, V-net: fully convolutional neural networks for volumetric medical image segmentation, Proc. - 2016 4th Int. Conf. 3D Vision, 3DV 2016 (2016) 565–571, https://doi.org/10.1109/3DV.2016.79.

[25] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G.S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mane, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viegas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, X. Zheng, TensorFlow: large-scale machine learning on heterogeneous distributed systems, ArXiv Prepr. ArXiv 1603 (2016) 04467 (accessed August 23, 2019), http://arxiv.org/abs/1603.04467.

[26] D.P. Kingma, J. Ba, Adam: a method for stochastic optimization, ArXiv Prepr. ArXiv 1412 (2014) 6980 (accessed August 23, 2019), http://arxiv.org/abs/1412.6980.

[27] S. Chaganti, A. Balachandran, G. Chabin, S. Cohen, T. Flohr, B. Georgescu, P. Grenier, S. Grbic, S. Liu, F. Mellot, N. Murray, S. Nicolaou, W. Parker, T. Re, P. Sanelli, A.W. Sauter, Z. Xu, Y. Yoo, V. Ziebandt, D. Comaniciu, Quantification of tomographic patterns associated with COVID-19 from Chest CT, ArXiv (2020) (accessed October 15, 2020), http://www.ncbi.nlm.nih.gov/pubmed/32550252%0Ahttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC7280906.

[28] F. Shan, Y. Gao, J. Wang, W. Shi, N. Shi, M. Han, Z. Xue, Y. Shi, Lung Infection Quantification of COVID-19 in CT Images With Deep Learning, 2020 (accessed October 15, 2020), http://arxiv.org/abs/2003.04655.

[29] D.P. Fan, T. Zhou, G.P. Ji, Y. Zhou, G. Chen, H. Fu, J. Shen, L. Shao, Inf-Net: Automatic COVID-19 Lung Infection Segmentation from CT Images, IEEE Trans. Med. Imaging 39 (2020) 2626–2637, https://doi.org/10.1109/TMI.2020.2996645.

[30] G. Wang, X. Liu, C. Li, Z. Xu, J. Ruan, H. Zhu, T. Meng, K. Li, N. Huang, S. Zhang, A noise-robust framework for automatic segmentation of COVID-19 pneumonia lesions from CT images, IEEE trans, Med. Imaging. 39 (2020) 2653–2663, https://doi.org/10.1109/TMI.2020.3000314.

[31] D. Wu, K. Gong, C. Arru, F. Homayounieh, B. Bizzo, V. Buch, H. Ren, K. Kim, N. Neumark, W.Y. Tak, M.K. Kang, A. Carriero, L. Saba, I. Dayan, M. Masjedi, R. Babaei, M.K. Kalra, Q. Li, Severity and consolidation quantification of COVID-19 from CT images using deep learning based on hybrid weak labels, IEEE J. Biomed. Heal. Informatics. (2020) 1, https://doi.org/10.1109/JBHI.2020.3030224.

[32] Z. Guo, N. Guo, K. Gong, S. Zhong, Q. Li, Gross tumor volume segmentation for head and neck cancer radiotherapy using deep dense multi-modality network, Phys. Med. Biol. 64 (2019), https://doi.org/10.1088/1361-6560/ab440d.

[33] S. Vanderschueren, A. De Weerdt, M. Malbrain, D. Vankersschaever, E. Frans, A. Wilmer, H. Bobbaers, Thrombocytopenia and prognosis in intensive care, Crit. Care Med. 28 (2000) 1871–1876, https://doi.org/10.1097/00003246-200006000-00031.

[34] G. Lippi, M. Plebani, B.M. Henry, Thrombocytopenia is associated with severe coronavirus disease 2019 (COVID-19) infections: a meta-analysis, Clin. Chim. Acta 506 (2020) 145–148, https://doi.org/10.1016/j.cca.2020.03.022.

[35] F. Zhou, T. Yu, R. Du, G. Fan, Y. Liu, Z. Liu, J. Xiang, Y. Wang, B. Song, X. Gu, L. Guan, Y. Wei, H. Li, X. Wu, J. Xu, S. Tu, Y. Zhang, H. Chen, B. Cao, Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study, Lancet 395 (2020) 1054–1062, https://doi.org/10.1016/S0140-6736(20)30566-3.

[36] Y. Shi, X. Yu, H. Zhao, H. Wang, R. Zhao, J. Sheng, Host susceptibility to severe COVID-19 and establishment of a host risk score: findings of 487 cases outside Wuhan, Crit. Care 24 (2020) 108, https://doi.org/10.1186/s13054-020-2833-7.

[37] L. Tan, Q. Wang, D. Zhang, J. Ding, Q. Huang, Y.Q. Tang, Q. Wang, H. Miao, Lymphopenia predicts disease severity of COVID-19: a descriptive and predictive study, Signal Transduct. Target. Ther. 5 (2020), https://doi.org/10.1038/s41392-020-0148-4.

[38] L. Yan, H.-T. Zhang, J. Goncalves, Y. Xiao, M. Wang, Y. Guo, C. Sun, X. Tang, L. Jing, M. Zhang, X. Huang, Y. Xiao, H. Cao, Y. Chen, T. Ren, F. Wang, Y. Xiao, S. Huang, X. Tan, N. Huang, B. Jiao, C. Cheng, Y. Zhang, A. Luo, L. Mombaerts, J. Jin, Z. Cao, S. Li, H. Xu, Y. Yuan, An interpretable mortality prediction model for COVID-19 patients, Int. J. Mach. Intell. Sens. Signal Process. (2020) 1–6, https://doi.org/10.1038/s42256-020-0180-7.

[39] J. Gong, J. Ou, X. Qiu, Y. Jie, Y. Chen, L. Yuan, J. Cao, M. Tan, W. Xu, F. Zheng, Y. Shi, B. Hu, Multicenter development and validation of a novel risk nomogram for early prediction of severe 2019-Novel coronavirus pneumonia, SSRN Electron. J. (2020), https://doi.org/10.2139/ssrn.3551365.

[40] J. Lu, S. Hu, R. Fan, Z. Liu, X. Yin, Q. Wang, Q. Lv, Z. Cai, H. Li, Y. Hu, Y. Han, H. Hu, W. Gao, S. Feng, Q. Liu, H. Li, J. Sun, J. Peng, X. Yi, Z. Zhou, Y. Guo, J. Hou, ACP risk grade: a simple mortality index for patients with confirmed or suspected severe acute respiratory syndrome coronavirus 2 disease (COVID-19) during the early stage of outbreak in Wuhan, China, SSRN Electron. J. (2020), https://doi.org/10.2139/ssrn.3543603.

[41] J. Xie, D. Hungerford, H. Chen, S.T. Abrams, S. Li, G. Wang, Y. Wang, H. Kang, L. Bonnett, R. Zheng, X. Li, Z. Tong, B. Du, H. Qiu, C.-H. Toh, Development and External Validation of a Prognostic Multivariable Model on Admission for

Hospitalized Patients with COVID-19, SSRN Electron. J. (2020), https://doi.org/10.2139/ssrn.3562456.

[42] L. Yan, H.-T. Zhang, Y. Xiao, M. Wang, C. Sun, J. Liang, S. Li, M. Zhang, Y. Guo, Y. Xiao, X. Tang, H. Cao, X. Tan, N. Huang, B. Jiao, A. Luo, Z. Cao, H. Xu, Y. Yuan, Prediction of criticality in patients with severe Covid-19 infection using three clinical features: a machine learning-based prognostic model with clinical data in Wuhan, MedRxiv (2020), https://doi.org/10.1101/2020.02.27.20028027, 2020.02.27.20028027.

[43] K. Zhang, X. Liu, J. Shen, Z. Li, Y. Sang, X. Wu, Y. Zha, W. Liang, C. Wang, K. Wang, L. Ye, M. Gao, Z. Zhou, L. Li, J. Wang, Z. Yang, H. Cai, J. Xu, L. Yang, W. Cai, W. Xu, S. Wu, W. Zhang, S. Jiang, L. Zheng, X. Zhang, L. Wang, L. Lu, J. Li, H. Yin, W. Wang, O. Li, C. Zhang, L. Liang, T. Wu, R. Deng, K. Wei, Y. Zhou, T. Chen, J.Y. N. Lau, M. Fok, J. He, T. Lin, W. Li, G. Wang, Clinically applicable AI system for accurate diagnosis, quantitative measurements, and prognosis of COVID-19 pneumonia using computed tomography, Cell 181 (2020) 1423–1433, https://doi.org/10.1016/j.cell.2020.04.045, e11.